DOCUMENT RESUME

ED 092 632 95 TM 003 815

AUTHOR Bianchini, John C.; Loret, Peter G.
TITLE Anchor Test Study Supplement. Final Report. Volume 31, Project Report.
INSTITUTION Educational Testing Service, Berkeley, Calif.
SPONS AGENCY Office of Education (DHEW), Washington, D.C.
PUB DATE [74]
CONTRACT OEC-0-71-4758(248)
NOTE 47p.; For related documents, see TM003784-817

EDRS PRICE MF-$0.75 HC-$1.85 PLUS POSTAGE
DESCRIPTORS *Elementary School Students; *Equated Scores; Grade 4; Grade 5; Grade 6; Measurement Techniques; *Reading Tests; Standardized Tests; *Testing
IDENTIFIERS *Anchor Test Study; California Achievement Tests; Comprehensive Tests of Basic Skills; Gates MacGinitie Reading Test; Iowa Test of Basic Skills; Metropolitan Achievement Tests; Sequential Tests of Educational Progress; SRA Achievement Series; Stanford Achievement Tests

ABSTRACT

The Anchor Test Study provides a method for translating a pupil's score on any one of eight widely used standardized reading tests for Grades 4, 5, and 6 to a corresponding score on any of the other seven tests, as well as furnishing new nationally representative norms for each of the eight tests. In addition, the Study presents new estimates of alternate form reliability for each test, provides estimates of the intercorrelations among the tests, and explores empirically some methodological questions in test-equating. The major purpose of the Anchor Test Study Supplement is to equate the Gates MacGinitie Reading Test (1964) to the other seven tests in the Study: California Achievement Tests (1970), Comprehensive Tests of Basic Skills (1968), Iowa Test of Basic Skills (1971), Metropolitan Achievement Tests (1970), Sequential Tests of Educational Progress (1969), SRA Achievement Series (1971), and Stanford Achievement Tests (1964). This volume of the Supplement includes information on administration and processing, design of the sample, methodology, and results. Some materials originally included in this volume were eliminated due to copyright restrictions. (RC)

ED 092632

BEST COPY AVAILABLE

ANCHOR TEST STUDY SUPPLEMENT

FINAL REPORT

VOLUME XXXI

PROJECT REPORT

TM 003 815

November, 1973

EDUCATIONAL TESTING SERVICE
PRINCETON, NEW JERSEY

BERKELEY, CALIFORNIA

This report was prepared under Contract No. OEC-0-71-4758(248) and submitted to the National Center for Educational Statistics, Office of Education, U.S. Department of Health, Education, and Welfare. The contractor was encouraged to exercise professional judgment in its contents; therefore, this report does not necessarily reflect positions and policies of the Government. In particular, no official endorsement of the tests mentioned in the report should be inferred.

TABLE OF CONTENTS

Page

CHAPTER 1 THE ANCHOR TEST STUDY SUPPLEMENT

1. The Nature of the Study 1
2. The Need for the Study 1
3. Design and Conduct of the Study 2

CHAPTER 2 ADMINISTRATION AND PROCESSING

1. The Invitational Phase 3
   1.1 State Invitations 3
   1.2 District Invitations 4
   1.3 School Invitations 5
   1.4 Participation 7
2. Materials Preparation and Shipping 7
   2.1 Packing, Shipping and Control 9
3. Material Receipt/Scoring and Reporting 13
   3.1 Receiving 14
   3.2 Scoring and Reporting 15

CHAPTER 3 DESIGN OF THE SAMPLE

1. Introduction 17
2. Determination of the Sample Size Used for the Supplemental Study 18
3. Basic Design Used for the Supplemental Study 19
4. Detailed Sample - Selection Procedures for the Supplemental Study 22
   4.1 The Selection of 56 of the School-Units Contained in the Primary Anchor Test Equating Sample 22
   4.2 The Selection of the Sample for the Supplemental Study 23
5. Testing - Sequence Assignments, Weighting, and Variance Estimation 25

CHAPTER 4 METHODOLOGY

1. Equating Study 26
   1.1 Weighting Procedure 26
   1.2 Equating Procedures 27
   1.3 Estimation of Equating Error 31
   1.4 Correlations 32
   1.5 Conditional Error of Equating 32
2. Development of Individual Norms and School Mean Norms for the GMT 33

CHAPTER 5 RESULTS

1. Equating Study 35
   1.1 Equating Sample 35
   1.2 Comparability of GMT and MAT Tests 37
   1.3 Equating Results 37
   1.4 Equating Error 38
2. Anchor Test Study Norms for the GMT 40

LIST OF TABLES

		Page
3-1	Comparison of the Allocations of the Anchor Test Equating Sample and Supplemental Sample to SES Groups.....	21
4-1	Reference Table for the Equating Results	30
4-2	Reference Table for Estimates of Equating Error	32
5-1	Anchor Test Equating Study Supplement: Equating Sample - Number of Pupils by Racial Groups and IQ Groups with Usable or Absentee, Incomplete, or Invalid Answer Sheets	36
5-2	Anchor Test Equating Study Supplement: Grades 4, 5 and 6 - Estimates of Averaged Error in Equating the GMT (64 ED.) Reading Test to the MAT (70 ED.) Reading Test (Average of Procedures 1 and 2, Equipercentile).....	39

CONTENTS OF ANCHOR TEST STUDY

SUPPLEMENT REPORT

1st Volume	Project Report
2nd Volume	Equating Tables
3rd Volume	Error of Equating Correlations Conditional Error of Equating Procedure Comparison Graphs Individual Norms and School Mean Norms Norms Comparisons

CHAPTER 1

THE ANCHOR TEST STUDY - SUPPLEMENT

1. The Nature of the Study

The Anchor Test Study supplement had as its major purpose the equating of the Gates-MacGinitie Reading Test to the seven standardized reading tests included in the original Anchor Test Study* and the provision of new nationally representative norms for this test. Tables for these purposes have been prepared for 4th grade, 5th grade, and 6th grade children.

This Gates-MacGinitie equating supplement involved the administration of pairs of tests to 4th, 5th, and 6th graders in a sample of public and nonpublic schools. Two appropriate reading tests (Form F of the Metropolitan Reading Test and Form D1 of the Gates-MacGinitie test or Forms D1 and D2 of the Gates-MacGinitie test) were administered to participating pupils in order to derive equivalent scores. The results of this supplementary study and those of the original Anchor Test Restandardization Study were combined to produce new norms for the Gates-MacGinitie.

2. The Need for the Study

The Anchor Test Study supplement permits the merging of data on the Gates-MacGinitie Reading Test with data obtained from the original Anchor Test Study. School personnel may convert Gates-MacGinitie Reading Test scores to those of the other seven tests, and vice versa.

---

*See Final Report, Anchor Test Study, Project Report and Volumes I-XXX, Educational Testing Service, December, 1972.

The test editions used in this supplement were the latest available from the publishers. The levels of tests used were those recommended by the test publishers. The test editions, forms, and levels used are listed below.

Title/Edition/Form	Publisher	Subtests	LEVEL USED FOR GRADE: 4	5	6
Gates-MacGinitie Reading Tests Survey DM1M, DM2M	Teachers College Press	Vocabulary Comprehension	Survey DM	Survey DM	Survey DM
Metropolitan Reading Tests (1970), Form F	Harcourt Brace Jovanovich	Word Analysis Reading	Elementary	Intermediate	Intermediate

3. Design and Conduct of the Study

The U.S. Office of Education developed the basic plan and detailed specifications for the original Anchor Test Study. The design for this supplement was developed by Educational Testing Service in cooperation with Westat Research, Inc. and the U.S. Office of Education. The contract for conduct of the Anchor Test Study Supplement, as well as the original Study, was awarded to Educational Testing Service (ETS), a nonprofit, nonstock corporation. Both the original study and the supplement were conducted by ETS's Berkeley, California office.

CHAPTER 2

ADMINISTRATION AND PROCESSING*

1. The Invitational Phase

The invitational phase of the Anchor Test Study Supplement required communications with the responsible individuals at three levels: Chief State School Officer, District (Local Education Agency) Superintendent, and School Principal. As in the original study, specifications required that approval be received from a higher administrative level prior to the issuance of an invitacion to any potential participant, i.e., it was necessary to obtain approval from the State level before issuing any District (LEA) Superintendent invitation. Similarly, it was necessary to receive District Superintendent approval prior to the issuance of school invitations.

1.1 State Invitations

Beginning in the week of February 19, 1973, telephone contact was made with CEIS (Committee on Evaluation and Information Systems) members and State Coordinators previously appointed for the original Anchor Test Study. Only those states containing primary sample schools for the supplement were contacted at this time. These individuals were informed of the purpose of the supplement, given the names of selected districts and schools in their state, and were asked for approval to proceed with district contacts. All individuals so contacted gave their approval, and provided the names and

---

*Samples of all documents, forms, procedures, and materials that are identical to those used in the original Anchor Test Study may be found in Volumes XXVIII and XXIX of the Final Report. Those items unique to the Supplement have been included in this report.

addresses of the appropriate district superintendents.

1.2 District Invitations

Upon approval at the state level and receipt of the District Superintendents' names and addresses, it was possible to begin the district invitation process. Since the beginning of this phase depended on contacts and approvals from the State Representatives, the district invitation phase varied from state to state, beginning in February 1973 and lasting into April 1973.

The written district invitations consisted of four components: a letter of invitation, a list of selected schools within the district, four Anchor Test Study brochures with a supplement description inserted, and a reply postcard. The letter of invitation was reviewed and approved by USOE and signed by Absalom Simms, Acting Director, Division of Intergovernmental Statistics. Written invitations were released only after an initial telephone contact had been made with the district office.

District Superintendents were asked to approve the participation of selected schools in their districts and to appoint a District Coordinator. The District Coordinator would verify the existence of selected schools (as well as the existence of grades 4, 5, and 6), provide the names and addresses of school principals, help with the distribution of school invitations, and assist in the resolution of any communication problems. A list of selected schools was transmitted to the Superintendent. Four Anchor Test Study brochures were supplied, together with an insert describing the Supplement, so that appropriate personnel could be informed about the Study. A return reply postcard was enclosed for the District Superintendent's response.

The reply card provided for the name, address, and phone number of the District Coordinator appointee. It also provided for the option of a later response to the invitation, and/or a request for more information. Most of the District invitations were released during late February and March, although some new back-up District invitations were released as late as April. A return reply control system was maintained so that non-respondents could be contacted and appropriate back-up districts invited. State Coordinators were kept informed of the status of invitations for the districts in their state.

1.3 School Invitations

All school invitation packets, addressed to individual principals, were mailed to the responsible individual at the district level together with the written district invitation. This permitted direct distribution through the District office, and school principals immediately became aware that the District had approved participation. It also served to introduce the individual appointed as District Coordinator. A covering memorandum, sent with these packets, requested that the District Coordinator distribute the invitations to the schools, help in describing the project to the principals, and assure that the school(s) completed and returned the reply forms contained in the invitation packet.

The invitation letter requested that the school principal approve participation for his/her school and appoint an individual to serve as School Coordinator. This individual would have primary responsibility for the actual test administration and data collection. The School Coordinator would complete a Pre-Test Information Form, be responsible for proper test

administration, receive and distribute test materials, and complete the school and class summary forms. Since these tasks would involve other than class time, the School Coordinator would receive an honorarium based on the number of pupils tested:

Honorarium for Test Coordination

149 or fewer pupils - $30

150 to 249 pupils - $40

250 or more pupils - $50

In addition to the letter of invitation, selected schools received four items in their invitation packet. A copy of the brochure entitled "A Description of the Anchor Test Study", together with an insert describing the Supplement, was enclosed. A reply postcard was provided for the name, address, and phone number of the School Coordinator. The card also provided for indications that the school was still considering the request and/or wished to have additional information. The postcard was completed by the principal and given to the School Coordinator with the remaining two items in the packet, the Pre-Test Information Form and a business reply envelope. Agreement to participate was indicated by the return of the completed postcard and the Pre-Test Information Form.

School invitation packets were first released in February and mailings continued until April. Schools unable to accept the invitation were replaced by schools designated as back-ups in the sample. Depending upon the characteristics of the originally selected school, the specific back-ups were located either in the same district, or in another district and/or state. In some cases, this required going back one step in the process and issuing an invitation to a District Superintendent not previously invited.

During the entire school invitation process, the District Coordinators,

State Coordinators, and the USOE were kept informed of the status of all school invitations. Telephone follow-ups with District Coordinators and principals were made to elicit rapid responses. In many cases, School Coordinator assignments were made by the principal during the telephone follow-up, with the promise that a completed Pre-Test Information Form would be forthcoming.

1.4 Participation

This Final Report contains a summary of invitational responses. This summary was prepared by testing units. It includes: the number of schools selected in the original Westat sample, the number of these schools found to be no longer in existence, schools ineligible because of test duplication(s), primary sample schools unable to accept invitations to participate, and substitute schools added. Seventy-one percent (71%) of the schools in the original sample agreed to participate. Of the total schools agreeing to participate (original sample schools and substitutes), all schools (100%) actually administered tests in the Anchor Test Study Supplement and returned answer sheets and ancillary documents to Educational Testing Service.

2. Materials Preparation and Shipping

School participation in the Anchor Test Study Supplement was confirmed by the return of an agreement postcard and a completed Pre-Test Information Form. The postcard designated a School Coordinator, the contact for all subsequent Study transactions. The Pre-Test Information Form served three primary purposes: it updated the school description file on the USOE computer tapes, it supplied additional input data required for the Study, and it

provided the information needed to assign tests and prepare shipments. More specifically, the Pre-Test Information Form contained the following information:

(a) A verification of the school address, and a confirmation of the School Coordinator's name and telephone number.

(b) A selected day for testing. Participating schools were required to contact Educational Testing Service if testing was not possible during the specified administration week (April 23-27, 1973).

(c) An indication that one or both of the Anchor Test Study Supplement instruments were already a part of the school's testing program. This assured that participants would not be assigned tests that duplicated in-school administrations.

(d) A listing of all classes at grades 4, 5, and 6. Each class was identified by a teacher name, and information was provided concerning the existence of split sessions, the number of pupils, physically handicapped or mentally retarded students, and a categorization of ability grouping.

(e) Answers to three "school description" questions, including an indication of the degree of urbanization of the community served by the school, the general income level of the school-family population, and the socio-economic level of the school area.

Pre-Test Information Forms were received and recorded in a master control system, assuring complete returns from all participants. The forms were verified to assure the existence of two classes each at grades 4, 5, and 6 and/or to assure that the grades represented were the appropriate component of a combined-schools testing unit (i.e., "pseudo-school"). All

Pre-Test Information Forms were thoroughly edited, and School Coordinators were contacted for any missing or incomplete information.

Processing of the Pre-Test Information Forms required the designation of the two tests to be administered. According to procedures outlined by Westat Research, Inc., schools were designated for test-pair assignment. Any duplication, between the tests assigned and tests already administered as part of the school's own program, was resolved, if possible, by re-assignment or substitution. For the majority of cases, this was not a factor, since possible duplications had already been determined during contact with the school district. These test assignments, together with the class descriptions, provided the basic information necessary for the computer preparation of control forms and shipping orders.

2.1 Packing, Shipping, and Control

For each participating school, a series of data processing forms for packing, shipping, and control were computer-produced. The Shipping Notice served as the master control form. Produced for each individual school, it listed the School Coordinator and Test Administrator kits (materials) to be packaged and shipped. The School Coordinator was provided with a kit for each separate test form and level being administered at the school including:

(a) A School Coordinator's Manual

(b) A School Coordinator's Report Form

(c) A 5% overage of test books and answer sheets

(d) Copies of the Test Administrator's Manual

(e) Business reply envelopes for the return of answer sheets

(f) Notification postcards (acknowledging receipt of the shipment and the completion of testing)

The Test Administrator's kit provided the materials necessary to administer one test to a class or group of students. It included:

(a) A Test Administrator's Manual

(b) A Class Identification Form

(c) Test books and answer sheets

The Shipping Notice served several functions. It was used by Educational Testing Service to package and prepare the materials for the School Coordinator and Test Administrator kits. For each Coordinator/Administrator entry listed on the Shipping Notice, the specified materials were collected and packaged as a unit. The Shipping Notice assured that the correct number of kits (for the appropriate publisher's tests) were prepared for each school. One copy of the Shipping Notice was inserted in a see-through enclosure and attached to the shipping carton. This served as the mailing address for the shipment, as well as providing a receiving check-list for the School Coordinator. The Coordinators were asked to count and verify that the shipment included all the materials necessary for the test administration.

In order to identify the various individual kits within a shipment, a computer-produced label was printed to match the entries on the Shipping Notice. A label was attached to each kit, identifying it as containing a specified test for a particular teacher. Because two tests were administered in the Supplement, each teacher received two personally addressed kits. While both kits were addressed to the same teacher, the labels indicated that different tests were enclosed.

Since alternate classes, within a school, were asked to administer tests in an alternate sequence, it was necessary to clearly identify the test to be administered first and the test to be administered second.

Therefore, a second label was affixed to each kit. It indicated whether that kit of test materials was the first or second test to be administered to the class. In order to verify that the correct labels had been applied to each kit, the kits were carefully cross-checked against the sequence of entries on the Shipping Notice.

In addition to the Shipping Notice, each School Coordinator received a computer-produced document entitled "School Coordinator's Report Form". This form was designed to provide additional testing information and to help the Coordinator organize the return of completed answer sheets. It listed all the class/teacher titles and tests, individually, and suggested that the Coordinator use the form to verify complete returns. School Coordinators were asked to indicate the grade level for each participating class group, the number of students in the class actually tested, and the total number of pupils enrolled in the class. The information was requested twice: once for each of the two tests administered. Participants were also asked to indicate the actual sequence of their test administrations, as well as the one test for which they preferred to have scores reported. The back of the School Coordinator's Report Form was used to record any over-all school irregularities during the test administration.

While the School Coordinator's Report Form served as a cover document for the entire school, a computer-produced "Class Identification Form" served as the cover document for each separate class group. This Class Identification Form was included with the materials in each Test Administrator's Kit. It was completed by the Test Administrator, relayed to the School Coordinator, and returned with the completed answer sheets. The Class Identification Form's primary function was the identification of a group of answer sheets as being those of a particular class. It

also served as a reminder to the Test Administrator to follow all instructions. The Test Administrator completed the Class Identification Form by recording the number of accompanying answer sheets and then returning both the answer sheets and the Form to the School Coordinator. He also recorded the sequence of each administration, i.e., whether the test was administered first or second. The back of the Class Identification Form served as a record of class or individual testing irregularities.

Each School Coordinator kit contained a 5% overage of test books and answer sheets for a particular test form and level. An extra copy of the Test Administrator's Manual was included in the kit, as well as a special manual for the School Coordinator. Instructions for the actual administration of the test, as well as Test Administrator procedures for recording information about each participating pupil, may be found in this Final Report.

Two return reply postcards were included in each school shipment. The School Coordinator was instructed to mail the first postcard to ETS upon receipt of the shipment. It served as a notification that the test materials had been delivered and that the school was supplied with all the materials necessary for testing. A control log was established at Educational Testing Service and all schools that had not forwarded their "Acknowledgement of Receipt" postcard, within approximately three weeks of the shipping date, were contacted by telephone. Most follow-up calls indicated that the shipments had been delivered, but that the cartons had not been opened (the card therefore not returned). Duplicate materials were prepared and shipped when there was the possibility of a shipment lost in transit. The School

Coordinator mailed the second postcard at the same time the completed answer sheets and cover documents were released. This postcard permitted the establishment of a follow-up system for return shipments that might have been lost or delayed in the mail.

3. Material Receipt/Scoring and Reporting

The week of April 23rd was designated as the primary week for schools to administer the Anchor Test Study instruments. The weeks of April 16th and April 30th were established as alternate administration weeks. Approximately 80% of the participants were able to administer the tests during the primary testing week. The alternate weeks were used by those schools with program conflicts.

A master listing of all participating schools was maintained at ETS and, as completed answer sheets and cover documents were returned, the listing was updated to show the current status of each individual school. This listing indicated the receipt of the "answer sheets mailed" postcard, the partial return of test materials, the complete return of answer sheets and cover documents, and the scoring/reporting unit (cycle number) in which the school would be processed. This system provided an instant status report for each participating school and identified those schools with problems and/or delays. Through immediate telephone follow-up, it was possible to obtain completed answer sheets and cover documents from all (100%) of those schools that had agreed to participate.

A status report was prepared and released to District Coordinators on May 16, 1973.

3.1 Receiving

The initial step in material-return receiving was a verification that basic cover and identification documents had been returned. Incoming packages were scanned to ascertain that each school had enclosed the School Coordinator's Report Form (SCRF) and a Class Identification Form (CIF) for each group of answer sheets. When these documents had been identified, the completed "answer sheet groups" were matched with group identification entries listed on the SCRF and any incomplete schools were set aside to await additional shipments. Irregularity reports were scanned to determine situations that might require special handling.

When it was established that two sets of answer sheets had been enclosed for each participating class and that no irregularities existed, the answer sheets were forwarded for further scanning and editing. This more detailed editing involved a verification of the completeness and accuracy of the school's input data. The School Coordinator's Report Form, the Class Identification Form, and the answer sheets were carefully reviewed and any overall discrepancies were resolved according to the rules specified in the original "Anchor Test Editing Procedures". More detailed discrepancies in counts were verified and adjusted as specified in the original "Anchor Test Study Receiving Rules". Individual answer sheets were numbered and a machine readable "Header Sheet" was coded for class/group identification. The "Anchor Receiving Log" was prepared, as a master control, by logging the date of answer sheet receipt, the School Coordinator Report Form number, the school's identification number, the identification number of each class in the school, and the number of answer sheets received for each class.

3.2 Scoring and Reporting

Optical transcription of header sheet and answer sheet information was the initial data processing step in the scoring and reporting cycle. Information was transcribed directly to magnetic tape by the NCS optical transcription machine. Computer editing programs scanned the transcribed data and produced a listing of all student and/or school information that appeared to be in conflict. This information was corrected, updated, and re-introduced into the system by re-processing the original input documents. All editing specifications outlined in the "Anchor Test Editing Procedures" were implemented in order to assure "clean" and valid data. During the course of all processing, stringent quality control procedures were maintained in order to assure accuracy. These procedures are outlined in the report entitled "Quality Control Procedures for The Anchor Test Study", in the original Anchor Test Study Final Report.

While the data for all students and schools were being processed and matched to the master computer file, for use in subsequent analyses, basic score reports were provided for immediate release to the participants. Based on each publisher's current normative data, a roster of results was prepared for each school. Since each pupil in the Supplement had taken two tests, the school was allowed to select the one test for which it preferred to have scores reported.

Class score rosters provided raw scores, the publisher's national percentile rank, and a stanine score for each pupil tested. For each class, the number and percentage of pupils falling within each quartile of the publisher's national norms was shown. The report also provided the class "N" (number tested), the raw score mean, and the raw score standard deviation.

In addition to the class data for each school, a combined grade level summary was prepared. This grade level report showed the raw score mean, the standard deviation, the pupil "N", and the number and percent of students falling in each quartile of the publisher's national norms (for all classes in that grade combined). In order to help schools understand the scores they were receiving, all reports included a description entitled "Interpreting Score Reports" as in the original Study. This document served to define the scores and normative information, as well as explaining the use of summary data.

Roster preparation and release was scheduled to provide results to all schools as quickly as possible, hopefully prior to the end of the school year. All school reports were released to participants by the first week in June. The reports for tardy respondents and schools with irregularities were mailed during subsequent weeks. By June 15th, all school reports had been released and all School Coordinators had been sent their honoraria for assisting with the Anchor Test Study Supplement. Upon completion of score reporting, the data anaylsis was begun.

CHAPTER 3

DESIGN OF THE SAMPLE

1. Introduction

As described in the introduction to Chapter 3 of the Anchor Test Study Report (p.27), the primary purposes of the study were to provide new national norms for pupils in grades 4, 5, and 6, and to establish equating relationships between scores on the seven reading tests used. In the present study, the Gates-MacGinitie Test (GMT) has been added to the original Anchor Test Study results. That is, national norms for the GMT and equating relationships between the GMT and each of the original seven tests have been established.

An attempt was made to design the supplemental sample in a way that would provide GMT norms and equating relationships with precision comparable to that obtained in the Anchor Test Study. This was done, to a large extent, by paralleling the design of the supplemental study with that of the Anchor Test Study. However, since it was not feasible to include all seven tests used in the Anchor Test Study, only the GMT and the "anchor" test (i.e., the MAT) were administered in the supplemental study. This simplified the design considerably compared to that used in the Anchor Test Equating Study, which involved the assignment of 21 different pairs of tests to the selected school-units.*

---

*As indicated on p. 36 of the Anchor Test Study Report, a school-unit consists of either an individual school or a "pseudo" school (i.e., a group of schools which had to be formed so that all school-units in the sampling frame for the Equating Study would contain at least two classes in each of grades 4, 5, and 6, as required by the testing design).

Because of the design simplification indicated above, it was not possible to predict precisely the magnitude of the standard errors to be obtained in the supplemental study on the basis of those found in the Anchor Test Study. However, as described later, the sample design allowed for the computation of standard errors of equating in the same way they were computed for the Anchor Test Equating Study. Furthermore, the sample design was such that it was expected to yield standard errors comparable to those obtained for the seven tests originally included in the Anchor Test Equating Study.

The standard errors of equating obtained in the supplemental study are given in Table 5-2, and, as expected, they are generally comparable to those reported in Tables 5-15, 5-16, and 5-17 in Chapter 5 of the Anchor Test Study Report.

2. Determination of the Sample Size Used for the Supplemental Study

A sample of 48 school-units was selected for joint administration of the GMT and the MAT. Another eight school-units were selected for test-retest administration of the GMT. To evaluate the sample size of 48 in the context of the Anchor Test Equating Study, two equating sample sizes must be noted. First, in the Anchor Test Equating Study, a given test combination (e.g., MAT and CAT) was administered in only 16 school-units (i.e., once in each of 16 sets of 28 school-units*), which is only one-third of the MAT-GMT sample size. Second, in the Anchor Test Equating Study, a given test was administered with other tests in a total of 96 school-units

*The description of the formation of these 16 sets of school-units is given in the Anchor Test Study Report on p. 36 and pp. 96-102.

(i.e., six in each of the 16 sets), which is twice as many as the 48 selected for this supplementary study. However, because of the substantially higher number of school-units that administered the specific test combination (i.e., GMT-MAT), and because of the use of basically the same stratification variables (as discussed below), it was reasonable to expect that the standard errors would be about the same for the GMT-MAT equating as for the equating of any test to the MAT in the Anchor Test Equating Study.

The sample of eight school-units for GMT test-retest administration was only half as many as that used for a given test in the Anchor Test Equating Study. However, it was felt that eight school-units would be adequate to estimate the parallel forms reliability of the GMT. (The use of 16 school-units in the Anchor Test Equating Study for test-retest administrations was actually a consequence of the equating design used, involving 16 groups, rather than a choice made specifically for estimating reliability.) The eight schools selected for the test-retest administration were not used in computing the equating relationships between the GMT and MAT.

3. Basic Design Used for the Supplemental Study

The list of school-units available for the sampling frame of the supplemental study was the set of all school-units selected for either the primary sample or for back-ups (i.e., possible substitutes) for the Anchor Test Equating Study.* From this frame, a sample was drawn for the supplemental study which had characteristics similar to those of the sample for the Anchor Test Equating Study.

---

*At the time the supplemental sample was selected, the sampling frame used for the Anchor Test Equating Study was no longer available.

The initial* sample of 48 school-units for the MAT-GMT administration in the supplemental study consisted of three first-level substitutes of primary selections (or in some cases the primary selections themselves) from each of the 16 groups of 28 school-units originally selected for the Anchor Test Equating Study. This provided six school-units from each of the eight principal strata.** Also, one additional first-level substitute (or primary selection) was selected from each principal stratum to yield eight school-units for the test-retest administration. Since the eight principal strata were defined using the SES ordering of major strata, the sample of 56 school-units for the supplemental study spanned essentially the same SES range as that for the Anchor Test equating sample. Furthermore, the double sampling rate for the high- and low-SES groups was preserved, and the proportion of the sample contained in each of the SES groups was approximately the same as it was for the Anchor Test Equating Study as demonstrated in Table 3-1.

---

*To avoid confusion in referring to the "primary" equating sample (i.e., the sample as initially selected before any substitutions were made) and the "primary" sample for the supplemental study, the primary sample for the supplemental study will be referred to as the "initial" sample.

**Principal strata are defined in the Anchor Test Study Report on pages 99-101. A description of the selection of the primary sample and the substitutes is given in Sections 6.3 and 6.9 of Chapter 3 of that report.

Table 3-1. Comparison of the Allocations of the Anchor Test Equating Sample and Supplemental Sample to SES Groups

SES group*	Approximate number of school-units selected for the Anchor Test Equating Study (percent)	Number of school-units selected for Supplemental Study (percent)
High	112 (25)	14 (25)
Medium	224 (50)	28 (50)
Low	112 (25)	14 (25)
	448 (100)	56 (100)

*The determination of SES groups is described in Section 6.2 of Chapter 3 of the Anchor Test Study Report.

Use of second- or higher-level equating substitutes as the initial sample for the supplemental study was not possible because of the requirement for the Anchor Test Equating Study that these substitutes not be from the same districts as the corresponding primary school-units. A limited investigation was carried out which indicated that the use of second-level substitutes would have led to an underrepresentation of schools located in the larger districts. Since district size may be correlated to some extent with test scores, second- or higher-level substitutes were not used for the initial sample for the supplemental study. However, the use of only first-level substitutes as the initial sample would have underrepresented schools in school districts that do not contain more than one school-unit within a major stratum.

Thus, to avoid underrepresenting schools in either larger or smaller districts, the following basic sampling method was used. First, a stratified sample of seven of the primary school-units was selected from each of the eight principal strata for the Anchor Test Equating Study, as described in detail in the next section. Whenever one of these 56 primary equating school-units contained a first-level substitute, that substitute was taken into the initial sample for the supplemental study. If a selected primary school-unit did not contain a first-level substitute, the primary unit itself was taken into the initial sample. The resulting sample provided the appropriate representation of schools from both the larger and smaller districts.

In the implementation of the basic design, some minor modifications were required. These modifications and the procedures used for substituting for initial selections that refused to participate are described in the next section.

4. Detailed Sample-Selection Procedures for the Supplemental Study

4.1 The Selection of 56 of the School-Units Contained in the Primary Anchor Test Equating Sample

Based entirely on the ordering of the final strata in the Anchor Test equating sample (described on p. 101 of the Anchor Test Study Report), the 28 primary school-units in each of the 16 sets were grouped into four strata of seven school-units each. One school-unit was then selected at random from each of these four strata. Also, one of the four school-units selected from each set was randomly designated for GMT test-retest administration. This produced eight school-units from each principal stratum for a total of 64, including 48 allocated to the GMT-MAT assignment and 16 allocated

to the GMT test-retest assignment. The 16 test-retest assignments consisted of two school-units from each of the eight principal strata. This group was reduced to eight school-units by randomly discarding one of the two school-units selected from each principal stratum. Thus, a sample of 56 primary equating school-units from the Anchor Test Equating Study was used as a base to obtain the initial sample of 56 school-units for the supplemental study.

The Selection of the Sample for the Supplemental Study

From the 56 primary equating school-units selected by the procedure described above, the initial sample of 56 school-units and necessary substitutes for the supplemental study were obtained using the following steps:

(1) Primary equating school-unit that had a first-level substitute:

(a) If the primary school-unit had participated in the Anchor Test Equating Study, the first-level substitute was taken into the initial sample. If the first-level substitute school-unit could not participate in the supplemental study, then the second-, third-, fourth-, and fifth-level substitutes served as possible substitutes for this school-unit.

(b) If the primary school-unit had not participated in the Anchor Test Equating Study, and the first-level substitute had, the primary school-unit was taken into the initial sample. If the primary school-unit declined to participate in the supplemental study, the second-, third-, fourth-, and fifth-level substitutes served as possible substitutes for this school-unit.

(c) If neither the primary nor the first-level substitute school-unit had participated in the Anchor Test Equating Study, the primary school-unit was taken into the initial sample. If a district

refusal was received, then the higher-level substitutes served as possible substitutes for the initial school-unit. If a district acceptance was received and the primary school-unit could not participate, participation was requested from the first-level substitute before proceeding to the higher-level substitutes. In neither case was the higher-level substitute that had participated in the Anchor Test Equating Study included as a possible substitute for the initial school-unit.

(2) Primary school-unit that did not have a first-level substitute:

(a) If the primary school-unit had participated in the Anchor Test Equating Study, it was taken into the initial sample unless it had administered the MAT in the Anchor Test Equating Study. If the school-unit had administered the MAT, a substitute was obtained from among the other 27 primary equating school-units in the same group. The substitute was taken as the nearest school-unit on the list of 28 which did not have a first-level substitute and had not administered the MAT in the Anchor Test Equating Study. Whenever possible, the substitute was taken to be a school-unit in the same major stratum as the first selection. If the substitute could not participate, then further substitutes were obtained from the second-, third-, fourth-, and fifth-level substitutes for the primary equating school-unit that was first selected.

(b) If the primary school-unit had not participated in the Anchor Test Equating Study, the primary school-unit was taken into the initial sample. If this school-unit could not participate in the supplemental study, the second-, third-, fourth-, and

fifth-level back-up school-units were used to obtain a substitute. However, in such a case, the substitute which had participated in the Anchor Test Equating Study was not reinvited for the supplemental study.

A few cases arose which were not covered adequately by the above general guidelines. In these cases the above guidelines were extended to provide a selection procedure.

5. Testing-Sequence Assignments, Weighting, and Variance Estimation

Once the sample of 56 cooperating school-units was obtained, the testing of pupils was carried out in a way similar to that for the Anchor Test Equating Study. That is, the classes in each school-unit were numbered, and the odd-numbered classes were randomly assigned to a testing sequence. The opposite testing sequence was used for the even-numbered classes.

The weights assigned to the pupils in each participating school-unit were computed in the same way as in the Anchor Test Equating Study.* The same variables and principles were used to form the pupil-nonresponse weighting classes, although of course, fewer pupils were available in this case.

An eight-degrees-of-freedom estimate of the standard error of equating the MAT to the GMT was computed using the same balanced half-sample replication method as was used in the Anchor Test Equating Study. In this case, the half-samples consisted of 24 school-units -- three from each of the 8 principal strata -- rather than 224 as in the Anchor Test Equating Study. However, the estimate of the variance of an equated score was computed as it was for the Anchor Test Equating Study.

---

*See Section 9 of Chapter 3 of the Anchor Test Study Report. It should be noted that the probability of selection from the Anchor Test Equating Study sample was constant for all school-units selected for the supplemental study; thus, no adjustment of school weight was required.

CHAPTER 4

METHODOLOGY

1. Equating Study

The test administration design for equating the Gates-MacGinitie Reading Test (GMT) to each of the seven reading tests originally included in the Anchor Test Study required, at each of grades 4, 5, and 6, a counterbalanced administration of the GMT and the MAT, as well as a counterbalanced administration of two alternate forms of the GMT for estimating parallel forms reliability. The procedures for the selection and the assignment of school-units to each of the two test-pairs are described in Sections 4 and 5 of Chapter 3. Within a given school-unit, half of the classes at each grade were tested in one order of administration and half in the opposite order. The administration of test-pairs in a given sequence yielded six scores for each student tested: vocabulary, reading comprehension, and total reading, for both the test taken first and the test taken second.

1.1 Weighting Procedure

For equating purposes, each student was assigned a weight which was thereafter associated with his test scores whenever they were included in any of the data analyses. The weight consisted of the product of three* factors determined in the same manner as in the Anchor Test Equating Study. (Weighting rationale and computational formulas are given in Section 9 of Chapter 3 of the Anchor Test Study Report.)

---

*Factor 4, an adjustment to render equal N's for the 56 different test-pair/order of administration combinations in the Anchor Test Equating Study, was not included in the weighting procedure for this study because it was irrelevant.

Factor 1 was a school weight determined as the inverse of the school selection probability. Adjustment for non-participating schools was made within major strata as described in Section 1.1 of Chapter 4 in the Anchor Test Study Report.

Factor 2 was an adjustment for absenteeism and invalid test data. It was determined by membership in a particular "weighting class" defined by sex, race, and IQ level within each major stratum (details of which are also given in Chapter 4, Section 1.1 in the Anchor Test Study Report).

Factor 3 was an adjustment for equalizing the probability that a class would be assigned to each of the two orders of administration (details of which are given in Section 2.1 of Chapter 4 in the Anchor Test Study Report).

1.2 Equating Procedures

The objective of the equating study was to equate the scores on each GMT subtest and total reading test* to the corresponding subtest scores and total reading scores of each of the seven reading tests originally included in the Anchor Test Study in grades 4, 5, and 6. Thus, the equating study consisted of nine analogues of the equating process, all treated identically with respect to the equating operations. Each application of the equating process produced equating tables between corresponding pairs of subtests or total tests for each of the two equating procedures described below.

For each of the nine analogues, the equating process consisted of (a) the frequency estimation procedure (details of which are presented in

*For this study the GMT total reading score was defined as the sum of the Vocabulary and Comprehension scores. The Speed and Accuracy section of the test was not included.

Appendix J of the Anchor Test Study Report), and (b) the equipercentile method of equating (details of which are given in Appendix K of the Anchor Test Study Report) applied through each of two procedures.

Procedure 1 utilized the data from the MAT when given as a second test in the GMT-MAT administration and the corresponding MAT data base from the Anchor Test Equating Study to estimate, for the entire Anchor Test Study equating sample, the marginal distribution of the GMT when taken as a first test. Specifically, a bivariate distribution with GMT as the ordinate and MAT as the abscissa was generated from the data of the GMT-MAT order of administration. Then the MAT base distribution used to estimate the final marginal distributions in Procedure 1 of the Anchor Test Study (see Operational Equating Trials in Section 2.3 of Chapter 4 in the Anchor Test Study Report) was used to estimate the GMT marginal distribution. The estimated GMT marginal distribution was then smoothed analytically using the Tukey-Cureton procedure (details of which are given in Appendix G of the Anchor Test Study Report). This smoothed distribution, paired with each of the corresponding smoothed distributions from Procedure 1 of the Anchor Test Study, was used to equate the GMT to each of the original seven tests by means of the equipercentile method.

Procedure 2, conversely, utilized the data from the MAT when given as a first test in the MAT-GMT order of administration and the corresponding MAT data base from the Anchor Test Equating Study to estimate the GMT distribution when taken as a second test. Thus, for Procedure 2, the frequency estimation procedure was applied analogously, with the roles of first and second tests reversed. The resulting estimated GMT distribution

was smoothed analytically and paired with the corresponding smoothed distributions obtained by Procedure 2 in the Anchor Test Study, for equating the GMT to each of the seven original tests by the equipercentile method.

The final equating tables were then obtained by averaging the equating results of Procedures 1 and 2 so that the order of administration factor explicit in both equating procedures would be averaged.

The equating tables produced by Procedures 1 and 2, and the final equating tables which were generated by averaging the equating results of Procedures 1 and 2, are reported in the Anchor Test Study Supplement. These tables appear in the order indicated by the Reference Table 4-1. It should be noted that these tables contain the equating results obtained by the same procedures in the Anchor Test Study for each of the original seven tests as well as the equating results for the GMT obtained in the supplemental study.

TABLE 4-1. Reference Table for the Equating Results

GRADE	TEST	TABLE NUMBERS		
		Procedure 1	Procedure 2	Average of Proc. 1 and 2
4	Vocabulary	2188 - 2195	2196 - 2203	2204 - 2211
	Reading Comprehension	2212 - 2219	2220 - 2227	2228 - 2235
	Total Reading	2236 - 2243	2244 - 2251	2252 - 2259
5	Vocabulary	2260 - 2267	2268 - 2275	2276 - 2283
	Reading Comprehension	2284 - 2291	2292 - 2299	2300 - 2307
	Total Reading	2308 - 2315	2316 - 2323	2324 - 2331
6	Vocabulary	2332 - 2339	2340 - 2347	2348 - 2355
	Reading Comprehension	2356 - 2363	2364 - 2371	2372 - 2379
	Total Reading	2380 - 2387	2388 - 2395	2396 - 2403

1.3 Estimation of Equating Error

In order to estimate empirically the error of equating associated with the MAT equivalent scores for the GMT, the equating operations (for Procedures 1 and 2, equipercentile method) used to equate the GMT to the MAT were replicated eight times on random half-samples of schools. The assignment of school-units to each of the eight half-samples was carried out as in the Anchor Test Study. In this case, however, each half-sample consisted of eight groups of 3 school-units each, not groups of 28 school-units as in the Anchor Test Equating Study. (Details of assignment of school-units to each of the eight half-samples are given in Section 11 of Chapter 3 in the Anchor Test Study Report.)

For each replication, the MAT data base that was used to estimate the final marginal distributions of the original seven tests for the corresponding replication in the Anchor Test Study was used to estimate the GMT marginal distributions in this study. The resulting GMT marginal distribution was then paired with the appropriate MAT distribution obtained from the same replication in the Anchor Test Study. Then, using these pairs of distributions, the GMT was equated to the MAT exactly as was done in the actual equating of the two tests in the supplemental study. Thus, the eight replications yielded eight sets of equating tables for each of the three procedures, i.e., Procedure 1, Procedure 2, and the Average of Procedures 1 and 2. The error in equating the GMT to the MAT was calculated in the same way as in the Anchor Test Equating Study (see Section 2.11 of Chapter 4 in the Anchor Test Study Report). The results of equating error estimates for the three procedures are given in Tables 2313 through 2321, as indicated in the reference table below.

TABLE 4-2. Reference Table for Estimates of Equating Error

TEST	TABLE NUMBERS		
	Procedure 1	Procedure 2	Average of Proc. 1 and 2
GMT Vocabulary	2313	2316	2319
GMT Reading Comprehension	2314	2317	2320
GMT Total Reading	2315	2318	2321

1.4 Correlations

For each of the two orders of administration of the GMT and MAT, and of alternate forms of the GMT, correlations were computed in the same way as in the Anchor Test Study for corresponding pairs of vocabulary, reading comprehension, and total reading scores at each of grades 4, 5, and 6. (Computational formulas are given in Appendix L of the Anchor Test Study Report.)

The observed correlations, both unaveraged and averaged, and correlations corrected for attenuation are given in Tables 2322 through 2325.

1.5 Conditional Error of Equating

The conditional error of equating was calculated for the MAT scores when transformed to the GMT equivalent scores obtained by the Average of Procedures 1 and 2 only. The test score data for the two orders of administration were combined for the computation of the conditional error and were about equally divided between first and second administrations. The procedure for calculating the conditional error was the same as that used in the Anchor Test Equating Study (see Section 2.12 of Chapter 4 in the Anchor Test Study Report).

The MAT conditional error of equating for vocabulary, reading comprehension, and total reading, for each of grades 4, 5, and 6, is shown in Tables 2326 through 2328.

2. Development of Individual Norms and School Mean Norms for the GMT

The individual norms for the GMT were obtained by using the MAT equivalent scores for the GMT obtained in this study and the MAT individual norms obtained in the Anchor Test Restandardization Study. Specifically, each raw score point on the GMT was mapped into the corresponding MAT mid-percentile rank and stanine through its MAT equivalent score obtained by the Average of Procedures 1 and 2, in exactly the same way as in the Anchor Test Study. (For details, see Section 1.3 of Chapter 4 in the Anchor Test Study Report.)

The GMT individual norms for each of the three scores, i.e., vocabulary, reading comprehension, and total reading, and for each of grades 4, 5, and 6, are given in Tables 2304 through 2306.

The school mean norms for the GMT were obtained by using the GMT equivalent score for the MAT obtained in this study and the MAT raw score data from the Anchor Test Restandardization Study. The MAT score for each individual student in the Anchor Test Study restandardization sample was transformed to its corresponding GMT equivalent score obtained by the Average of Procedures 1 and 2, and then the school means for the GMT were computed on the transformed scores, i.e., on the GMT equivalent scores.* The school mean norms for the GMT were then developed

---

*For computation of the school means, the individual student scores carried the original Anchor Test Restandardization Study weights (see Section 1.1 of Chapter 4 in the Anchor Test Study Report).

in the same way as in the Anchor Test Restandardization Study. (For details, see Section 1.5 of Chapter 4 in the Anchor Test Study Report.) The GMT school mean norms for each of the three reading scores, and for each of grades 4, 5, and 6, are given in Tables 2307 through 2309.

CHAPTER 5

RESULTS

1. Equating Study

For the supplemental equating study, the GMT and the MAT were administered to a sample of pupils at each of grades 4, 5, and 6, during the period of April 23 to April 27, 1973. These data, together with the Anchor Test Study data collected in the spring of 1972, made possible the subsequent development of equivalency tables between the GMT and the seven tests originally included in the Anchor Test Study. Additionally, these data made possible the extension of the Anchor Test Study norms to the GMT through the use of the GMT-MAT equivalent scores and the MAT normative data obtained in the Anchor Test Study.

1.1 Equating Sample

The equating design required a total of 56 school-units, 48 of which were assigned to the GMT-MAT test administration, and 8 to the GMT alternate forms administration. The obtained sample consisted of 80 individual schools which as a group comprised 47 school-units for the GMT-MAT test administration, (one school-unit was lost at the last minute because of a conflicting test administration requirement) and 8 school-units for the GMT parallel forms administration.

Table 5-1 shows the numbers of children that participated in the study for the three grades combined. The counts are given in a race x IQ bivariate table showing the number of usable, incomplete, invalid, and absentee answer sheets in each cell. It is noteworthy that 93% of the children enrolled in the 80 schools provided usable test data, which compares favorably with the 94% obtained in the Anchor Test Study.

TABLE 5-1. ANCHOR TEST EQUATING STUDY SUPPLEMENT: EQUATING SAMPLE

Number of Pupils by Racial Groups and IQ Groups with Usable or Absentee, Incomplete, or Invalid Answer Sheets

IQ	Type of Answer Sheet	RACIAL GROUPS: American Indian	Black	Oriental	Spanish Surname	White or Other	No Information	Total Number	Total Percentage*
Below 75	Usable	1	64	2	17	68	2	154	90.1
	Absentee		3		1	6		10	
	Incomplete		5		1	1		7	
	Invalid								
75 - 89	Usable	21	309	4	90	758	3	1185	93.0
	Absentee	3	16		9	31		59	
	Incomplete		10		2	15		27	
	Invalid					3		3	
90 - 110	Usable	39	498	29	225	3841	14	4646	94.8
	Absentee	4	15		10	138		167	
	Incomplete		23		7	56		86	
	Invalid					4		4	
111 - 125	Usable	14	83	20	45	2479	10	2651	95.3
	Absentee	1	1		4	69	2	77	
	Incomplete		2		3	47		52	
	Invalid					2		2	
Above 125	Usable	3	9	5	10	704		731	95.4
	Absentee		1		1	23		25	
	Incomplete					9		9	
	Invalid					1		1	
No Information	Usable	17	1343	23	446	2847	360	5036	90.0
	Absentee	3	59	1	70	131	84	348	
	Incomplete		76	3	24	70	31	204	
	Invalid		1			5		6	
Total	Usable	95	2306	83	833	10697	389	14403	93.0
	Absentee	11	95	1	95	398	86	686	
	Incomplete		116	3	37	198	31	385	
	Invalid		1			15		16	
Percentage*		86.4	91.6	95.4	86.3	94.6	76.9		

* The percentage of usable answer sheets for each marginal cell (i.e., racial group or IQ level).

1.2 Comparability of GMT and MAT Tests

As discussed in the Anchor Test Study Report (see Section 2.2 of Chapter 5), correlations corrected for the unreliability of the tests of .95 or higher are desirable when equating non-parallel tests. In the present study, all the correlations between GMT and MAT for corresponding subtests and total test, which are reported in Tables 2324 and 2325, are higher than .95. Thus these correlations support the inference that the two reading tests measure essentially the same intellectual functions and reading skills, and that their equating is justified.

1.3 Equating Results

The results of the Anchor Test Supplemental Study are reported in Tables 2188 through 2403. The first three sets of 24 tables each are the equating tables for vocabulary, reading comprehension, and total reading scores in that order for grade 4. Similarly, the next three sets are for grade 5, and the last three sets are for grade 6. Each of the nine sets of 24 tables each is preceded by a list (accessed by a red-lettered tab) indicating the order of appearance of the tables in that set. The within-set table sequence is the same for all three grades, presenting first the equating tables produced by Procedure 1, followed by those produced by Procedure 2, and those produced by the Average of Procedures 1 and 2.

It should be noted that, in order to facilitate the use of the equivalent scores obtained for the seven tests originally included in the Anchor Test Study and those obtained in the supplemental study, these equating tables contain both the equating results of the Anchor Test Study

and those of the supplemental study. Specifically, the first seven columns of each equating table for the CAT, CTBS, ITBS, MAT, STEP II, SRA, and SAT contain the same data as the corresponding equating tables presented in the Anchor Test Study Report. The eighth column (i.e., the GMT column) contains the equating results obtained in the supplemental study, as does each of the columns of the GMT equating table.

The relationships between equating lines obtained by Procedure 1 and the corresponding ones obtained by Procedure 2 are presented in graphs 6049 through 6174. For each grade the set of graphs is presented in the same way. Each set is preceded by a list (accessed by a red-lettered tab) showing the abbreviations, symbolisms, cross-references with equating table, and the order in which the graphs appear. Each set consists of 24 graphs (eight each for the vocabulary, reading comprehension, and total reading) showing the equating lines between the GMT and each of the other seven tests.

As was the case in the Anchor Test Study, the graphs show that the results obtained by Procedures 1 and 2 were indeed consistent, and thus the averaging of the corresponding equatings was justified.

1.4 Equating Error

The estimated error of equating for the GMT when equated to the MAT is reported in Tables 2313 through 2321. The equating error estimates are presented separately for the vocabulary, reading comprehension, and total reading tests for each grade and equating procedure.

While the equating error estimates for Procedures 1 and 2 provide an additional basis for evaluating all the equating results, only those associated with the Average of Procedures 1 and 2 are of practical importance,

TABLE 5-2. ANCHOR TEST EQUATING STUDY SUPPLEMENT: GRADES 4, 5 and 6

Estimates of Averaged* Error in Equating the GMT (64 ED.) Reading Test to the MAT (70 ED.) Reading Test

Average of Procedures 1 and 2, Equipercentile

Raw Score Range	GMT (64 ED.) Survey D - Form 1 GRADE 4 Average Error	GMT (64 ED.) Survey D - Form 1 GRADE 5 Average Error	GMT (46 ED.) Survey D - Form 1 GRADE 6 Average Error
Vocabulary			
46 - 50	.1	.1	.1
41 - 45	.1	.2	.2
36 - 40	.1	.2	.3
31 - 35	.1	.2	.3
26 - 30	.3	.3	.3
21 - 25	.3	.3	.3
16 - 20	.7	.3	.3
11 - 15	.5	.5	.4
6 - 10	.4	.7	.5
1 - 5	.4	1.4	.7
Reading Comprehension			
46 & above	.1	.2	.2
41 - 45	.2	.3	.3
36 - 40	.2	.3	.3
31 - 35	.3	.3	.3
26 - 30	.3	.3	.3
21 - 25	.3	.3	.3
16 - 20	.3	.3	.2
11 - 15	.4	.3	.2
6 - 10	.5	.4	.4
1 - 5	.7	.8	1.1
Total Reading			
91 & above	.2	.6	.3
81 - 90	.2	.3	.3
71 - 80	.2	.4	.4
61 - 70	.2	.4	.4
51 - 60	.3	.4	.6
41 - 50	.4	.4	.5
31 - 40	.5	.4	.3
21 - 30	.9	.7	.3
11 - 20	.9	1.6	1.0
1 - 10	1.6	1.2	2.6

* For vocabulary and reading comprehension, error estimates were averaged over intervals of five raw score points and smoothed. For the total reading test, estimates were averaged over intervals of ten and then smoothed.

since the equivalency tables derived in this particular manner were the ones used in developing the GMT norms and will be those included in the forthcoming User's Manual.

Table 5-2 contains a summarized version of the estimated error of equating for the Average of Procedures 1 and 2. The error associated with individual score points was averaged and smoothed to reduce the slight fluctuations at adjacent raw score points and to facilitate the use of these estimates in subsequent equivalent score interpretations. The estimates in this table indicate that the equating error associated with any GMT score above the chance range is but a fraction of a raw score point and is generally comparable to the estimated error for each of the tests originally included in the Anchor Test Study (see Section 2.5 of Chapter 5 in the Anchor Test Study Report).

2. Anchor Test Study Norms for the GMT

The individual norms and school mean norms for the GMT reading scores are shown in Tables 2304 through 2309. These norms were derived by applying the equating results obtained by the Average of Procedures 1 and 2 to the Anchor Test Study MAT norms, as was done for the six non-MAT tests in the Anchor Test Study. Because of the small equating error estimates, the consistency of equating results across procedures, and the high correlations between corresponding reading scores on the two tests, the norms provided for the GMT may be generally considered as reliable as those of the MAT itself.

In Tables 2310 through 2312 are shown the Anchor Test Study individual norms and the corresponding publisher's norms for the vocabulary and reading comprehension subtests. These data may be summarized as follows.

For the vocabulary subtest, the two sets of norms are generally comparable, with the largest difference being three percentile rank points. For the reading comprehension, however, the differences between the two sets of norms are both larger and systematic, especially within the mid-range at grade 4, where the Anchor Test Study norms yield consistently higher percentile ranks up to a maximum difference of 10 points. The same consistent trend is also present at grade 5, but the magnitude of the differences is somewhat less than that observed at grade 4. At grade 6, the two sets of reading comprehension norms are more in line than those at grades 4 and 5, although the Anchor Test Study norms still yield slightly higher percentile ranks in the range of the 20th through the 70th percentile rank. Thus, on the basis of these comparative data, the users of the GMT reading comprehension subtest at grades 4 and 5 should expect somewhat higher percentile ranks with the use of the Anchor Test Study norms.